

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Yu, Guo-Liang
Rosen, Craig
- (ii) TITLE OF INVENTION: Colon Specific Genes and Proteins
- (iii) NUMBER OF SEQUENCES: 24
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi,
Stewart & Olstein
(B) STREET: 6 Becker Farm Road
(C) CITY: Roseland
(D) STATE: NJ
(E) COUNTRY: USA
(F) ZIP: 07068-1739
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/469,667
(B) FILING DATE: 06-JUN-1995
(C) CLASSIFICATION:
- (vii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Ferraro, Gregory D.
(B) REGISTRATION NUMBER: 36,134
(C) REFERENCE/DOCKET NUMBER: 325800-435
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 201-994-1700
(B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 638 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 1..501

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..501

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCC AGG CAG CTG GCT GCC SAC CAG GCC GTG TAT GTG AAG GTC AAG GCT	48
Ala Arg Gln Leu Ala Ala Xaa Gln Ala Val Tyr Val Lys Val Lys Ala	
1 5 10 15	
GAA GCC CGG GAA CTG CTG GGC CAC CCG TGG TCT CTG TGT CCT GTC TGT	96
Glu Ala Arg Glu Leu Leu Gly His Pro Trp Ser Leu Cys Pro Val Cys	
20 25 30	
GGG TGC CAA CTC ACC ACC TTT GAT GGG GCC CGT GGT GCC ACC ACT CTC	144
Gly Cys Gln Leu Thr Thr Phe Asp Gly Ala Arg Gly Ala Thr Thr Leu	
35 40 45	
CTG GTG TCT ATG AAG CTC TCT TCC CGC TGC CCA GGA CTA CAG AAT ACC	192
Leu Val Ser Met Lys Leu Ser Ser Arg Cys Pro Gly Leu Gln Asn Thr	
50 55 60	
ATC CCC TGG TAC CGT GTA GTT GCC GAA GTC CAG ATC TGC CAT GGC AAA	240
Ile Pro Trp Tyr Arg Val Val Ala Glu Val Gln Ile Cys His Gly Lys	
65 70 75 80	
ACG GAG GCT GTG GGC CAG GTC CAC ATC TTC TTC CAG GAT GGG ATG GTG	288
Thr Glu Ala Val Gly Gln Val His Ile Phe Phe Gln Asp Gly Met Val	
85 90 95	
ACG TTG ACT CCA AAC AAG GGT GTG TGG GTG AAT GGT CTC CGA GTG GAT	336
Thr Leu Thr Pro Asn Lys Gly Val Trp Val Asn Gly Leu Arg Val Asp	
100 105 110	
CTC CCA GCT GAG AAG TTA GCA TCT GTG TCC GTG AGT CGT ACA CCT GAT	384
Leu Pro Ala Glu Lys Leu Ala Ser Val Ser Val Ser Arg Thr Pro Asp	
115 120 125	
GGC TCC CTG CTA GTC CGC CAG AAG GCA GGG GTC CAG GTG TGG CTT GGA	432
Gly Ser Leu Leu Val Arg Gln Lys Ala Gly Val Gln Val Trp Leu Gly	
130 135 140	
GCC AAT GGG AAG GTG GCT GTG ATT GTG AGC AAT GAC CAT GCT GGG AAA	480
Ala Asn Gly Lys Val Ala Val Ile Val Ser Asn Asp His Ala Gly Lys	
145 150 155 160	
CTG TGT GGG GGC CTK TGG AAA ATTTGACGGG GGACCAGACC AATGATTGGG	531
Leu Cys Gly Gly Xaa Trp Lys	
165	
ATGATTCCCA GGAGAAGCCA GCGATTGGGG AAWTGAGAG CGCAGGGACT TTCTYCCMCA	591
TGTTAATGGG CTGWTCCAG TTCATCCAC CAGGAACGAA GGATTTT	638

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 167 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala Arg Gln Leu Ala Ala Xaa Gln Ala Val Tyr Val Lys Val Lys Ala
 1 5 10 15
 Glu Ala Arg Glu Leu Leu Gly His Pro Trp Ser Leu Cys Pro Val Cys
 20 25 30
 Gly Cys Gln Leu Thr Thr Phe Asp Gly Ala Arg Gly Ala Thr Thr Leu
 35 40 45
 Leu Val Ser Met Lys Leu Ser Ser Arg Cys Pro Gly Leu Gln Asn Thr
 50 55 60
 Ile Pro Trp Tyr Arg Val Val Ala Glu Val Gln Ile Cys His Gly Lys
 65 70 75 80
 Thr Glu Ala Val Gly Gln Val His Ile Phe Phe Gln Asp Gly Met Val
 85 90 95
 Thr Leu Thr Pro Asn Lys Gly Val Trp Val Asn Gly Leu Arg Val Asp
 100 105 110
 Leu Pro Ala Glu Lys Leu Ala Ser Val Ser Val Ser Arg Thr Pro Asp
 115 120 125
 Gly Ser Leu Leu Val Arg Gln Lys Ala Gly Val Gln Val Trp Leu Gly
 130 135 140
 Ala Asn Gly Lys Val Ala Val Ile Val Ser Asn Asp His Ala Gly Lys
 145 150 155 160
 Leu Cys Gly Gly Xaa Trp Lys
 165

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 874 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..705

(ix) FEATURE:

(A) NAME/KEY: mat_peptide

(B) LOCATION: 1..705

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAG GAC TGC GTG TGC ACG GAC AAG GTG GAC AAC AAC ACC CTG CTC AAC	48
Gln Asp Cys Val Cys Thr Asp Lys Val Asp Asn Asn Thr Leu Leu Asn	
1 5 10 15	
GTC ATC GCC TGC ACC CAC GTG CCC TGC AAC ACC TCC TGC AGC CCT GGG	96
Val Ile Ala Cys Thr His Val Pro Cys Asn Thr Ser Cys Ser Pro Gly	
20 25 30	
TTC GAA CTC ATG GAG GCC CCC GGG GAG TGC TGT AAG AAG TGT GAA CAG	144
Phe Glu Leu Met Glu Ala Pro Gly Glu Cys Cys Lys Lys Cys Glu Gln	
35 40 45	
ACG CAC TGT ATC ATC AAA CGG CCC GAC AAC CAG CAC GTC ATC CTG AAG	192
Thr His Cys Ile Ile Lys Arg Pro Asp Asn Gln His Val Ile Leu Lys	
50 55 60	
CCC GGG GAC TTC AAG AGC GAC CCG AAG AAC AAC TGC ACA TTC TTC AGC	240
Pro Gly Asp Phe Lys Ser Asp Pro Lys Asn Asn Cys Thr Phe Phe Ser	
65 70 75 80	
TGC GTG AAG ATC CAC AAC CAG CTC ATC TCG TCC GTT TCC AAC ATC ACC	288
Cys Val Lys Ile His Asn Gln Leu Ile Ser Ser Val Ser Asn Ile Thr	
85 90 95	
TGC CCC AAC TTT GAT GCC AGC ATT TGC ATC CCG GGC TCC ATC ACA TTC	336
Cys Pro Asn Phe Asp Ala Ser Ile Cys Ile Pro Gly Ser Ile Thr Phe	
100 105 110	
ATG CCC AAT GGA TGC TGC AAG ACC TGC ACC CCT CGC AAT GAG ACC AGG	384
Met Pro Asn Gly Cys Cys Lys Thr Cys Thr Pro Arg Asn Glu Thr Arg	
115 120 125	
GTG CCC TGC TCC ACC GTC CCC GTC ACC ACG GAG GTT TCG TAC GCC GGC	432
Val Pro Cys Ser Thr Val Pro Val Thr Thr Glu Val Ser Tyr Ala Gly	
130 135 140	
TGC ACC AAG ACC GTC CTC ATG AAT CAT TGC TCC GGG TCC TGC GGG ACA	480
Cys Thr Lys Thr Val Leu Met Asn His Cys Ser Gly Ser Cys Gly Thr	
145 150 155 160	
TIT GTC ATG TAC TCG GCC AAG GCC CAG GCC CTG GAC CAC AGC TGC TCC	528
Phe Val Met Tyr Ser Ala Lys Ala Gln Ala Leu Asp His Ser Cys Ser	
165 170 175	
TGC TGC AAA GAG GAG AAA ACC AGC CAG CGT GAG GTG GTC CTG AGC TGC	576
Cys Cys Lys Glu Glu Lys Thr Ser Gln Arg Glu Val Val Leu Ser Cys	
180 185 190	
CCC AAT GGC GGC TCG CTG ACA CAC ACC TAC ACC CAC ATC GAG AGC TGC	624
Pro Asn Gly Gly Ser Leu Thr His Thr Tyr Thr His Ile Glu Ser Cys	
195 200 205	

CAG TGC CAG GAC ACC GTC TGC GGG CTC CCC ACC GGC ACC TCC CGC CGG	672
Gln Cys Gln Asp Thr Val Cys Gly Leu Pro Thr Gly Thr Ser Arg Arg	
210 215 220	
GCC CGG CGT TCC CCT AGG CAT CTG GGG AGC GGG TGAGCGGGGT GGGCACAGCC	725
Ala Arg Arg Ser Pro Arg His Leu Gly Ser Gly	
225 230 235	
CCTTCACTGC CCTCGACAGC TTTACCTCCC CCGGACCCTC TGAGCCTCCT AAGCTCGGCT	785
TCCTCTCTTC AGATATTAT TGTCTGAGTT TTTGTTCACT CCTTGCTTTC CAATAATAAA	845
CTCAGGGGGA CATGCAAAAA AAAAAAAAAA	874

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 235 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gln	Asp	Cys	Val	Cys	Thr	Asp	Lys	Val	Asp	Asn	Asn	Thr	Leu	Leu	Asn	1	5	10	15
Val	Ile	Ala	Cys	Thr	His	Val	Pro	Cys	Asn	Thr	Ser	Cys	Ser	Pro	Gly	20	25	30	
Phe	Glu	Leu	Met	Glu	Ala	Pro	Gly	Glu	Cys	Cys	Lys	Lys	Cys	Glu	Gln	35	40	45	
Thr	His	Cys	Ile	Ile	Lys	Arg	Pro	Asp	Asn	Gln	His	Val	Ile	Leu	Lys	50	55	60	
Pro	Gly	Asp	Phe	Lys	Ser	Asp	Pro	Lys	Asn	Asn	Cys	Thr	Phe	Phe	Ser	65	70	75	80
Cys	Val	Lys	Ile	His	Asn	Gln	Leu	Ile	Ser	Ser	Val	Ser	Asn	Ile	Thr	85	90	95	
Cys	Pro	Asn	Phe	Asp	Ala	Ser	Ile	Cys	Ile	Pro	Gly	Ser	Ile	Thr	Phe	100	105	110	
Met	Pro	Asn	Gly	Cys	Cys	Lys	Thr	Cys	Thr	Pro	Arg	Asn	Glu	Thr	Arg	115	120	125	
Val	Pro	Cys	Ser	Thr	Val	Pro	Val	Thr	Thr	Glu	Val	Ser	Tyr	Ala	Gly	130	135	140	
Cys	Thr	Lys	Thr	Val	Leu	Met	Asn	His	Cys	Ser	Gly	Ser	Cys	Gly	Thr	145	150	155	160
Phe	Val	Met	Tyr	Ser	Ala	Lys	Ala	Gln	Ala	Leu	Asp	His	Ser	Cys	Ser				

165

170

175

Cys Cys Lys Glu Glu Lys Thr Ser Gln Arg Glu Val Val Leu Ser Cys
180 185 190

Pro Asn Gly Gly Ser Leu Thr His Thr Tyr Thr His Ile Glu Ser Cys
195 200 205

Gln Cys Gln Asp Thr Val Cys Gly Leu Pro Thr Gly Thr Ser Arg Arg
210 215 220

Ala Arg Arg Ser Pro Arg His Leu Gly Ser Gly
225 230 235

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1209 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATTGGTGCTA	CCTGGCTCTC	CTGTCTCTGC	AGCTCTACAG	GTGAGGCCCA	GCAGAGGGAG	60
TTAGGGCTCGC	CATGTTTCTG	GTGAGCCAAT	TTGGCTGATC	TTGGGTGTCT	GAACAGCTAT	120
TTGGGTCCACC	CCAATCCCTT	TCAGCTGCTG	CTTAATGCCC	TGCTCTCTCC	CTGGCCCAACC	180
TTATAGAGAG	CCCAAGAGC	TCCTGTAAGA	GGGAGAATC	TATCTGTGGT	TTATAATCTT	240
GCACGAGGCA	CCAGAAGTCT	CCCTGGGTCT	TGTGAATGAA	CTACATTAT	CCCCTTTCCT	300
GCCCAACCA	CAAATCTTT	CCTTCAAAGA	GGCCTGCCT	GGTCCCCTCC	ACCCAATGC	360
ACCATGAGAT	CGGTCCAAGA	GTCCATTCCC	CAGGTGGGAG	CCAACTGTCA	GGGAGGTCTT	420
TCCCAACCAA	CATCTTTCAG	TTGCTGGGAG	GTGACCATAG	GGCTCTGCTT	TAAAGATAT	480
GGCTGCTTCA	AAGGCCAGAG	TCACAGGAAG	GACTTCTTCC	AGGGAGATTA	GTGGTGATGG	540
AGAGGAGAGT	TAAATGACC	TCATGTCCTT	CTGTGCCACG	GTTTGTGTGA	GTTTTCACTC	600
TTCTAATGCA	AGGGTCTCAC	ACTGTGAACC	ACTTAGGATG	TGATCACTTT	CAGGTGGCCA	660
GGAAATGTGA	ATGCTTTTGG	CTCAGTTCAT	CTAAAAAGA	TATCTATTTG	AAAGTTCTCA	720
GAGTTGTACA	TATGTTTCAG	AGTACAGGAT	CTGTACATAA	AAGTTCTTTT	CCTAAACCAT	780
TCACCAAGAG	CCAATATCTA	GGCATTTCCT	CGGTAGCACA	AATTTTCTNA	TTGCTTAGAA	840
AATTGTCTCT	CCTGTTCTTT	CTGTCTGNAG	ACTTAAGTGA	GTTAGGTCTT	TAAGGAAAGC	900

AACGCTCCTC TGAATGCTT GTCTTTTTC TGTGCCCAG ATAGCTGGTC CTTTTCGGG 960
 AGTTAGATGT ATAGAGTGT TGTATGTAAA CATTTCCTGT AGGCATCACC ATGAACANAG 1020
 ATATATTTTC TATTTANTTA NTATATGTGC ACTTCAAGAA GTCAGTGTCA GAGAAATAAA 1080
 GAATTGTCTT AAATGTCATG ATTGGAGATG TCCTTTGTCAT TGCTTGGAAG GGGTGTACCT 1140
 AGAGCCAAGG AAATTGGCTC TGGTTTGGAA AAATTTTGCT GTTATTATAG TAAACATACA 1200
 AAGGATGTC 1209

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 548 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..405

- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION: 1..405

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATG AGT CCT GTG AAA AAC AAT GTG GGC AGA GGC CTA AAC ATC GCC CTG 48
 Met Ser Pro Val Lys Asn Asn Val Gly Arg Gly Leu Asn Ile Ala Leu
 1 5 10 15

GTG AAT GGA ACC ACG GGA GCT GTG CTG GGA CAG AAG GCA TTT GAC ATG 96
 Val Asn Gly Thr Thr Gly Ala Val Leu Gly Gln Lys Ala Phe Asp Met
 20 25 30

TAC TCT GGA GAT GTT ATG CAC CTA GTG AAA TTC CTT AAA GAA ATT CCG 144
 Tyr Ser Gly Asp Val Met His Leu Val Lys Phe Leu Lys Glu Ile Pro
 35 40 45

GGG GGT GCA CTG GTG CTG GTG GCC TCC TAC GAC GAT CCA GGG ACC AAA 192
 Gly Gly Ala Leu Val Leu Val Ala Ser Tyr Asp Asp Pro Gly Thr Lys
 50 55 60

ATG AAC GAT GAA AGC AGG AAA CTC TTC TCT GAC TTG GGG AGT TCC TAC 240
 Met Asn Asp Glu Ser Arg Lys Leu Phe Ser Asp Leu Gly Ser Ser Tyr
 65 70 75 80

GCA AAA CAA CTG GGC TTC CGG GAC AGC TGG GTC TTC ATA GGA GCC AAA 288
 Ala Lys Gln Leu Gly Phe Arg Asp Ser Trp Val Phe Ile Gly Ala Lys

85

90

95

GAC CTC AGG GGT AAA AGC CCC TTT GAG CAG TTC TTA AAG AAC AGC CCA 336
 Asp Leu Arg Gly Lys Ser Pro Phe Glu Gln Phe Leu Lys Asn Ser Pro
 100 105 110

GAC ACA AAC AAA TAC GAG GGA TGG CCA GAG CTG CTG GAG ATG GAG GGC 384
 Asp Thr Asn Lys Tyr Glu Gly Trp Pro Glu Leu Leu Glu Met Glu Gly
 115 120 125

TGC ATG CCC CCG AAG CCA TTT TAGGGTGGCT GTGGCTCTTC CTCAGCCAGG 435
 Cys Met Pro Pro Lys Pro Phe
 130 135

GGCCTGAAGA AGYTCTGGCC TGCATTAGGA GTCANAGCCC GGCAGGCTGN AGGAGGAGGA 495

GCAGGGGGTG CTGCGTGGAA GGTGCTGCAG GCCTTGACAG CTGTGTCCGC CCT 548

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 135 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ser Pro Val Lys Asn Asn Val Gly Arg Gly Leu Asn Ile Ala Leu
 1 5 10 15

Val Asn Gly Thr Thr Gly Ala Val Leu Gly Gln Lys Ala Phe Asp Met
 20 25 30

Tyr Ser Gly Asp Val Met His Leu Val Lys Phe Leu Lys Glu Ile Pro
 35 40 45

Gly Gly Ala Leu Val Leu Val Ala Ser Tyr Asp Asp Pro Gly Thr Lys
 50 55 60

Met Asn Asp Glu Ser Arg Lys Leu Phe Ser Asp Leu Gly Ser Ser Tyr
 65 70 75 80

Ala Lys Gln Leu Gly Phe Arg Asp Ser Trp Val Phe Ile Gly Ala Lys
 85 90 95

Asp Leu Arg Gly Lys Ser Pro Phe Glu Gln Phe Leu Lys Asn Ser Pro
 100 105 110

Asp Thr Asn Lys Tyr Glu Gly Trp Pro Glu Leu Leu Glu Met Glu Gly
 115 120 125

Cys Met Pro Pro Lys Pro Phe
 130 135

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 878 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..685

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

T	GTC	TAC	TCA	AGG	TAT	TTC	ACA	ACT	TAT	GAC	ACG	AAT	GGT	AGA	TAC	46
Val	Tyr	Ser	Arg	Tyr	Phe	Thr	Thr	Tyr	Asp	Thr	Asn	Gly	Arg	Tyr		
1					5					10				15		
AGT	GTA	AAA	GTG	CGG	GCT	CTG	GGA	GGA	GTT	AAC	GCA	GCC	AGA	CGG	AGA	94
Ser	Val	Lys	Val	Arg	Ala	Leu	Gly	Gly	Val	Asn	Ala	Ala	Arg	Arg	Arg	
			20					25					30			
GTG	ATA	CCC	CAG	CAG	AGT	GGA	GCA	CTG	TAC	ATA	CCT	GGC	TGG	ATT	GAG	142
Val	Ile	Pro	Gln	Gln	Ser	Gly	Ala	Leu	Tyr	Ile	Pro	Gly	Trp	Ile	Glu	
			35				40						45			
AAT	GAT	GAA	ATA	CAA	TGG	AAT	CCA	AGA	AGA	CCT	GAA	ATT	AAT	AAG	GAT	190
Asn	Asp	Glu	Ile	Gln	Trp	Asn	Pro	Pro	Arg	Pro	Glu	Ile	Asn	Lys	Asp	
		50				55					60					
GAT	GTT	CAA	CAC	AAG	CAA	GTG	TGT	TTC	AGC	AGA	ACA	TCC	TCG	GGA	GGC	238
Asp	Val	Gln	His	Lys	Gln	Val	Cys	Phe	Ser	Arg	Thr	Ser	Ser	Gly	Gly	
		65			70					75						
TCA	TTT	GTG	GCT	TCT	GAT	GTC	CCA	AAT	GCT	CCC	ATA	CCT	GAT	CTC	TTC	286
Ser	Phe	Val	Ala	Ser	Asp	Val	Pro	Asn	Ala	Pro	Ile	Pro	Asp	Leu	Phe	
	80				85					90				95		
CCA	CCT	GGC	CAA	ATC	ACC	GAC	CTG	AAG	GCG	GAA	ATT	CAC	GGG	GGC	AGT	334
Pro	Pro	Gly	Gln	Ile	Thr	Asp	Leu	Lys	Ala	Glu	Ile	His	Gly	Gly	Ser	
			100					105					110			
CTC	ATT	AAT	CTG	ACT	TGG	ACA	GCT	CCT	GGG	GAT	GAT	TAT	GAC	CAT	GGA	382
Leu	Ile	Asn	Leu	Thr	Trp	Thr	Ala	Pro	Gly	Asp	Asp	Tyr	Asp	His	Gly	
		115					120					125				
ACA	GCT	CAC	AAG	TAT	ATC	ATT	CGA	ATA	AGT	ACA	AGT	ATT	CTT	GAT	CTC	430
Thr	Ala	His	Lys	Tyr	Ile	Ile	Arg	Ile	Ser	Thr	Ser	Ile	Leu	Asp	Leu	
		130				135					140					
AGA	GAC	AAG	TTC	AAT	GAA	TCT	CTT	CAA	GTG	AAT	ACT	ACT	GCT	CTC	ATC	478
Arg	Asp	Lys	Phe	Asn	Glu	Ser	Leu	Gln	Val	Asn	Thr	Thr	Ala	Leu	Ile	
	145					150				155						

CCA AAG GAA GCC AAC TCT GAG GAA GTC TTT TTG TTT AAA CCA GAA AAC 526
 Pro Lys Glu Ala Asn Ser Glu Glu Val Phe Leu Phe Lys Pro Glu Asn
 160 165 170 175

ATT ACT TTT GAA AAT GGC ACA GAT CTT TTC ATT GCT ATT CAG GCT GTT 574
 Ile Thr Phe Glu Asn Gly Thr Asp Leu Phe Ile Ala Ile Gln Ala Val
 180 185 190

GAT AAG GTC GAT CTG AAA TCA GAA ATA TCC AAC ATT GCA CGA GTA TCT 622
 Asp Lys Val Asp Leu Lys Ser Glu Ile Ser Asn Ile Ala Arg Val Ser
 195 200 205

TTG TTT ATT CCT CCA CAG ACT CCG CCA GAG ACA CCT AGT CCT GAT GAA 670
 Leu Phe Ile Pro Pro Gln Thr Pro Pro Glu Thr Pro Ser Pro Asp Glu
 210 215 220

ACG TCT GCT CCT TGT GCCTAATATT CATAACAACA GCACCATTC TGGCATTAC 725
 Thr Ser Ala Pro Cys
 225

CAATTTAAAAA TTATGTGGAA GTGGGTAGGA GAACTGCAGT TGTCATAGN CTAGGGGTGA 785
 CAATTTTGTGC GGTGAATAAA TAATSATTTT ANCTTTTTT TGRITTTATAA AAAAACGGNT 845
 NCCCATTTGGG NNTNTNGNGG GGGGGNNTTT TAA 878

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 228 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Val Tyr Ser Arg Tyr Phe Thr Thr Tyr Asp Thr Asn Gly Arg Tyr Ser
 1 5 10 15

Val Lys Val Arg Ala Leu Gly Gly Val Asn Ala Ala Arg Arg Arg Val
 20 25 30

Ile Pro Gln Gln Ser Gly Ala Leu Tyr Ile Pro Gly Trp Ile Glu Asn
 35 40 45

Asp Glu Ile Gln Trp Asn Pro Pro Arg Pro Glu Ile Asn Lys Asp Asp
 50 55 60

Val Gln His Lys Gln Val Cys Phe Ser Arg Thr Ser Ser Gly Gly Ser
 65 70 75 80

Phe Val Ala Ser Asp Val Pro Asn Ala Pro Ile Pro Asp Leu Phe Pro
 85 90 95

Pro Gly Gln Ile Thr Asp Leu Lys Ala Glu Ile His Gly Gly Ser Leu
 100 105 110

Ile Asn Leu Thr Trp Thr Ala Pro Gly Asp Asp Tyr Asp His Gly Thr
 115 120 125

Ala His Lys Tyr Ile Ile Arg Ile Ser Thr Ser Ile Leu Asp Leu Arg
 130 135 140

Asp Lys Phe Asn Glu Ser Leu Gln Val Asn Thr Thr Ala Leu Ile Pro
 145 150 155 160

Lys Glu Ala Asn Ser Glu Glu Val Phe Leu Phe Lys Pro Glu Asn Ile
 165 170 175

Thr Phe Glu Asn Gly Thr Asp Leu Phe Ile Ala Ile Gln Ala Val Asp
 180 185 190

Lys Val Asp Leu Lys Ser Glu Ile Ser Asn Ile Ala Arg Val Ser Leu
 195 200 205

Phe Ile Pro Pro Gln Thr Pro Pro Glu Thr Pro Ser Pro Asp Glu Thr
 210 215 220

Ser Ala Pro Cys
 225

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 560 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 2..490

- (ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION: 2..490

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

A GTC GCT CTC CTA GCC CTT CTC TGT GCC TCA CCC TCT GGC AAT GCC	46
Val Ala Leu Leu Ala Leu Leu Cys Ala Ser Pro Ser Gly Asn Ala	
1 5 10 15	
ATT CAG GCC AGG TCT TCC TCC TAT AGT GGA GAG TAT GGA GGT GGT GGT	94
Ile Gln Ala Arg Ser Ser Ser Tyr Ser Gly Glu Tyr Gly Gly Gly Gly	
20 25 30	

GGA AAG CGA TTC TCT CAT TCT GGC AAC CAG TTG GAC GGC CCC ATC ACC	142
Gly Lys Arg Phe Ser His Ser Gly Asn Gln Leu Asp Gly Pro Ile Thr	
35 40 45	
GCC CTC CGG GTC CGA GTC AAC ACA TAC TAC ATC GTA GGT CTT CAG GTG	190
Ala Leu Arg Val Arg Val Asn Thr Tyr Tyr Ile Val Gly Leu Gln Val	
50 55 60	
CGC TAT GGC AAG GTG TGG AGC GAC TAT GTG GGT GGT CGC AAC GGA GAC	238
Arg Tyr Gly Lys Val Trp Ser Asp Tyr Val Gly Gly Arg Asn Gly Asp	
65 70 75	
CTG GAG GAG ATC TTT CTG CAC CCT GGG GAA TCA GTG ATC CAG GTT TCT	286
Leu Glu Glu Ile Phe Leu His Pro Gly Glu Ser Val Ile Gln Val Ser	
80 85 90 95	
GGG AAG TAC AAG TGG TAC CTG AAG AAG CTG GTA TTT GTG ACA GAC AAG	334
Gly Lys Tyr Lys Trp Tyr Leu Lys Lys Leu Val Phe Val Thr Asp Lys	
100 105 110	
GGC CGC TAT CTG TCT TTT GGG AAA GAC AGT GGC ACA AGT TTC AAT GCC	382
Gly Arg Tyr Leu Ser Phe Gly Lys Asp Ser Gly Thr Ser Phe Asn Ala	
115 120 125	
GTC CCC TTG CAC CCC AAC ACC GTG CTC CGC TTC ATC AGT GGC CGG TCT	430
Val Pro Leu His Pro Asn Thr Val Leu Arg Phe Ile Ser Gly Arg Ser	
130 135 140	
GGT TCT CTC ATC GAT GCC ATT GGC CTG CAC TGG GAT GTT TAC CCC ACT	478
Gly Ser Leu Ile Asp Ala Ile Gly Leu His Trp Asp Val Tyr Pro Thr	
145 150 155	
AGC TGC AGC AGA TGCTGAGCCT CCTCTCCTTG GCAGGGGCAC TGTGATGAGG	530
Ser Cys Ser Arg	
160	
AGTAAGAAGT CCTTATCACT AACCCCATC	560

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Val Ala Leu Leu Ala Leu Leu Cys Ala Ser Pro Ser Gly Asn Ala Ile	
1 5 10 15	
Gln Ala Arg Ser Ser Tyr Ser Gly Glu Tyr Gly Gly Gly Gly	
20 25 30	
Lys Arg Phe Ser His Ser Gly Asn Gln Leu Asp Gly Pro Ile Thr Ala	
35 40 45	

Leu Arg Val Arg Val Asn Thr Tyr Tyr Ile Val Gly Leu Gln Val Arg
 50 55 60
 Tyr Gly Lys Val Trp Ser Asp Tyr Val Gly Gly Arg Asn Gly Asp Leu
 65 70 75 80
 Glu Glu Ile Phe Leu His Pro Gly Glu Ser Val Ile Gln Val Ser Gly
 85 90 95
 Lys Tyr Lys Trp Tyr Leu Lys Lys Leu Val Phe Val Thr Asp Lys Gly
 100 105 110
 Arg Tyr Leu Ser Phe Gly Lys Asp Ser Gly Thr Ser Phe Asn Ala Val
 115 120 125
 Pro Leu His Pro Asn Thr Val Leu Arg Phe Ile Ser Gly Arg Ser Gly
 130 135 140
 Ser Leu Ile Asp Ala Ile Gly Leu His Trp Asp Val Tyr Pro Thr Ser
 145 150 155 160
 Cys Ser Arg

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 709 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TAAACTTGCT GTTTGTTC TGTGTCTGT CTTGGTTGG TATTTTCAGTA AGTTTTTGGT	60
ATTCTCAAAT TTTATCTAA TGGATAAACT ATTAACATAG AACATAAACC CCAATTCCTC	120
ATTTTCATTT TCTCTTAGG ATGAATCATA CAAACTCAA TATAGAGCAA TGTTTGTAA	180
GAATGTGTT ATTAACAAAG AGGAGGTTCT AAGATATAAA GCCTCAGAGA ACAGGAAGAA	240
AAGCGGGTC CATAAGAAGA TGAGGTCTAA CCGGAAGAT GCTGCTGAGA AGGCAGAGAC	300
AGATGTGGA GAAATCTATC ACCCAGTCAT GTGCACTGAA TGTTCCACTG AAGTGGCAGT	360
TTACGACAAG GATGAAGTCT TTCATTTTTC CAATGTTTGA GCAAGCCATT CCTAACAGC	420
CCAACTGGCA TTTAATTACC CAATACTGTA TATAAGGCAA ATATGGACAG TTACTTTCCT	480
CTTGCTGTT CATATCCTTC AGTGACATTG AGGAAGCAGT GTTCTCTTT TAAAGGGGA	540

ATAGTTGTCA ACCITCATTC ATCTCTTACA TCTTCACCC TCTCCTTTT TTTTCTTTG	600
ATTTTCCCC TTATTGATGG GACTGATATT CATTCTGTTT TTGATGAACA TTTGGAAACT	660
GTCCGGGCTTT TTATTAAGC TCTGTAGAAT TAAATGTTT TGAATTAT	709

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 570 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 125..367

(ix) FEATURE:
 (A) NAME/KEY: mat_peptide
 (B) LOCATION: 125..367

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CAGGAGGGAG AGCCTTCCCC AAGCAAACAA TCCAGAGCAG CTGTGCAAAAC AACGGTGCAT	60
AAATAAGGCC TCCTGACCA TGAATGCGAG TCCGCTGAGC TGCGTACCGG AGCCCCACGGT	120
GGTC ATG GCT GCC AGA GCG CTC TGC ATG CTG GGG CTG GTC CTG GCC TTG	169
Met Ala Ala Arg Ala Leu Cys Met Leu Gly Leu Val Leu Ala Leu	
1 5 10 15	
CTG TCC TCC AGC TCT GCT GAG GAG TAC GTG GGC CTG TCT GCA AAC CAG	217
Leu Ser Ser Ser Ser Ala Glu Glu Tyr Val Gly Leu Ser Ala Asn Gln	
20 25 30	
TGT GCC GTG CCA GCC AAG GAC AGG GTG GAC TGC GGC TAC CCC CAT GTC	265
Cys Ala Val Pro Ala Lys Asp Arg Val Asp Cys Gly Tyr Pro His Val	
35 40 45	
ACC CCC AAG GAG TGC AAC AAC CGG GGC TGC TGC TTT GAC TCC AGG ATC	313
Thr Pro Lys Glu Cys Asn Asn Arg Gly Cys Cys Phe Asp Ser Arg Ile	
50 55 60	
CCT GGA GTG CCT TGG TGT TTC AAG CCC CTG ACA GGG AAG CAG GAA TGC	361
Pro Gly Val Pro Trp Cys Phe Lys Pro Leu Thr Gly Lys Gln Glu Cys	
65 70 75	
ACC TTC TGAGGCACCT CCAGCTGCCC CCCGCCGGG GGATGCGAGG CTCGGAGCAC	417
Thr Phe	
80	
CCTTGCCCGG CTGTGATTGC TGCCAGGCAC TGTTTCATCTC AGCTTTTCTG TCCCTTTGCT	477

CCCGGAAGCG CTTCTGCTGA AAGTTCATAT CTGGAGCCTG ATGTTTAACG TAGTCCCATG 537
CTCCACCCGA AAAAAAAAAA AAAAAAAAAA AAA 570

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Ala	Ala	Arg	Ala	Leu	Cys	Met	Leu	Gly	Leu	Val	Leu	Ala	Leu	Leu
1					5				10					15	
Ser	Ser	Ser	Ser	Ala	Glu	Glu	Tyr	Val	Gly	Leu	Ser	Ala	Asn	Gln	Cys
				20				25					30		
Ala	Val	Pro	Ala	Lys	Asp	Arg	Val	Asp	Cys	Gly	Tyr	Pro	His	Val	Thr
			35				40					45			
Pro	Lys	Glu	Cys	Asn	Asn	Arg	Gly	Cys	Cys	Phe	Asp	Ser	Arg	Ile	Pro
			50			55				60					
Gly	Val	Pro	Trp	Cys	Phe	Lys	Pro	Leu	Thr	Gly	Lys	Gln	Glu	Cys	Thr
			65			70				75				80	
Phe															

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1121 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 42..1010

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 42..1010

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAGCTCTTCT CACAGGACCA GCCACTAGCG CAGCTCGAGC G ATG GCC TAT GTC 53
Met Ala Tyr Val
1

CCC GCA CCG GGC TAC CAG CCC ACC TAC AAC CCG ACG CTG CCT TAC TAC Pro Ala Pro Gly Tyr Gln Pro Thr Tyr Asn Pro Thr Leu Pro Tyr Tyr 5 10 15 20	101
CAG CCC ATC CCG GGC GGG CTC AAC GTG GGA ATG TCT GTT TAC ATC CAA Gln Pro Ile Pro Gly Gly Leu Asn Val Gly Met Ser Val Tyr Ile Gln 25 30 35	149
GGA GTG GCC AGC GAG CAC ATG AAG CGG TTC TTC GTG AAC TTT GTG GTT Gly Val Ala Ser Glu His Met Lys Arg Phe Phe Val Asn Phe Val Val 40 45 50	197
GGG CAG GAT CCG GGC TCA GAC GTC GCC TTC CAC TTC AAT CCG CGG TTT Gly Gln Asp Pro Gly Ser Asp Val Ala Phe His Phe Asn Pro Arg Phe 55 60 65	245
GAC GGC TGG GAC AAG GTG GTC TTC AAC ACG TTG CAG GGC GGG AAG TGG Asp Gly Trp Asp Lys Val Val Phe Asn Thr Leu Gln Gly Gly Lys Trp 70 75 80	293
GGC GAG GAG GAG AGG AAG AGG AGC ATG CCC TTC AAA AAG GGT GCC GCC Gly Ser Glu Glu Arg Lys Arg Ser Met Pro Phe Lys Lys Gly Ala Ala 85 90 95 100	341
TTT GAG CTG GTC TTC ATA GTC CTG GCT GAG CAC TAC AAG GTG GTG GTA Phe Glu Leu Val Phe Ile Val Leu Ala Glu His Tyr Lys Val Val Val 105 110 115	389
AAT GGA AAT CCC TTC TAT GAG TAC GGG CAC CGG CTT CCC CTA CAG ATG Asn Gly Asn Pro Phe Tyr Glu Tyr Gly His Arg Leu Pro Leu Gln Met 120 125 130	437
GTC ACC CAC CTG CAA GTG GAT GGG GAT CTG CAA CTT CAA TCA ATC AAC Val Thr His Leu Gln Val Asp Gly Asp Leu Gln Leu Gln Ser Ile Asn 135 140 145	485
TTC ATC GGA GGC CAG CCC CTC CGG CCC CAG GGA CCC CCG ATG ATG CCA Phe Ile Gly Gly Gln Pro Leu Arg Pro Gln Gly Pro Pro Met Met Pro 150 155 160	533
CCT TAC CCT GGT CCC GGA CAT TGC CAT CAA CAG CTG AAC AGC CTG CCC Pro Tyr Pro Gly Pro Gly His Cys His Gln Gln Leu Asn Ser Leu Pro 165 170 175 180	581
ACC ATG GAA GGA CCC CCA ACC TTC AAC CCG CCT GTG CCA TAT TTC GGG Thr Met Glu Gly Pro Pro Thr Phe Asn Pro Pro Val Pro Tyr Phe Gly 185 190 195	629
AGG CTG CAA GGA GGG CTC ACA GCT CGA AGA ACC ATC ATC ATC AAG GGC Arg Leu Gln Gly Gly Leu Thr Ala Arg Arg Thr Ile Ile Lys Gly 200 205 210	677
TAT GTG CCT CCC ACA GGC AAG AGC TTT GCT ATC AAC TTC AAG GTG GGC Tyr Val Pro Pro Thr Gly Lys Ser Phe Ala Ile Asn Phe Lys Val Gly 215 220 225	725

TCC TCA GGG GAC ATA GCT CTG CAC ATT AAT CCC CGC ATG GGC AAC GGT	773
Ser Ser Gly Asp Ile Ala Leu His Ile Asn Pro Arg Met Gly Asn Gly	
230 235 240	
ACC GTG GTC CGG AAC AGC CTT CTG AAT GGC TCG TGG GGA TCC GAG GAG	821
Thr Val Val Arg Asn Ser Leu Leu Asn Gly Ser Trp Gly Ser Glu Glu	
245 250 255 260	
AAG AAG ATC ACC CAC AAC CCA TTT GGT CCC GGA CAG TTC TTT GAT CTG	869
Lys Lys Ile Thr His Asn Pro Phe Gly Pro Gly Gln Phe Phe Asp Leu	
265 270 275	
TCC ATT CGC TGT GGC TTG GAT CGC TTC AAG GTT TAC GCC AAT GGC CAG	917
Ser Ile Arg Cys Gly Leu Asp Arg Phe Lys Val Tyr Ala Asn Gly Gln	
280 285 290	
CAC CTC TTT GAC TTT GCC CAT CGC CTC TCG GCC TTC CAG AGG GTG GAC	965
His Leu Phe Asp Phe Ala His Arg Leu Ser Ala Phe Gln Arg Val Asp	
295 300 305	
AACA TTG GAA ATC CAG GGT GAT GTC ACC TTG TCC TAT GTC CAG ATC	1010
Thr Leu Glu Ile Gln Gly Asp Val Thr Leu Ser Tyr Val Gln Ile	
310 315 320	
TAATCTATTCT CTGGGGCCAT AACTCATGGG AAAACAGAAT TATCCCTAG GACTCCTTTC	1070
TAAGCCCTA ATAAATGTC TGAGGGTGTC TCAAAAAAAAA AAAAAAAAAA A	1121

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 323 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Ala Tyr Val Pro Ala Pro Gly Tyr Gln Pro Thr Tyr Asn Pro Thr	
1 5 10 15	
Leu Pro Tyr Tyr Gln Pro Ile Pro Gly Gly Leu Asn Val Gly Met Ser	
20 25 30	
Val Tyr Ile Gln Gly Val Ala Ser Glu His Met Lys Arg Phe Phe Val	
35 40 45	
Asn Phe Val Val Gly Gln Asp Pro Gly Ser Asp Val Ala Phe His Phe	
50 55 60	
Asn Pro Arg Phe Asp Gly Trp Asp Lys Val Val Phe Asn Thr Leu Gln	
65 70 75 80	
Gly Gly Lys Trp Gly Ser Glu Glu Arg Lys Arg Ser Met Pro Phe Lys	
85 90 95	

Lys Gly Ala Ala Phe Glu Leu Val Phe Ile Val Leu Ala Glu His Tyr
 100 105 110
 Lys Val Val Val Asn Gly Asn Pro Phe Tyr Glu Tyr Gly His Arg Leu
 115 120 125
 Pro Leu Gln Met Val Thr His Leu Gln Val Asp Gly Asp Leu Gln Leu
 130 135 140
 Gln Ser Ile Asn Phe Ile Gly Gly Gln Pro Leu Arg Pro Gln Gly Pro
 145 150 155 160
 Pro Met Met Pro Pro Tyr Pro Gly Pro Gly His Cys His Gln Gln Leu
 165 170 175
 Asn Ser Leu Pro Thr Met Glu Gly Pro Pro Thr Phe Asn Pro Pro Val
 180 185 190
 Pro Tyr Phe Gly Arg Leu Gln Gly Gly Leu Thr Ala Arg Arg Thr Ile
 195 200 205
 Ile Ile Lys Gly Tyr Val Pro Pro Thr Gly Lys Ser Phe Ala Ile Asn
 210 215 220
 Phe Lys Val Gly Ser Ser Gly Asp Ile Ala Leu His Ile Asn Pro Arg
 225 230 235 240
 Met Gly Asn Gly Thr Val Val Arg Asn Ser Leu Leu Asn Gly Ser Trp
 245 250 255
 Gly Ser Glu Glu Lys Lys Ile Thr His Asn Pro Phe Gly Pro Gly Gln
 260 265 270
 Phe Phe Asp Leu Ser Ile Arg Cys Gly Leu Asp Arg Phe Lys Val Tyr
 275 280 285
 Ala Asn Gly Gln His Leu Phe Asp Phe Ala His Arg Leu Ser Ala Phe
 290 295 300
 Gln Arg Val Asp Thr Leu Glu Ile Gln Gly Asp Val Thr Leu Ser Tyr
 305 310 315 320
 Val Gln Ile

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 605 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..603

(ix) FEATURE:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 1..603

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GTT GAT ATT AAA ACC AGT GAA ACC AAA CAT GAC ACC TCT CTG AAA CCT	48
Val Asp Ile Lys Thr Ser Glu Thr Lys His Asp Thr Ser Leu Lys Pro	
1 5 10 15	
ATT AGT GTC TCC TAC AAC CCA GCC ACA GCC AAA GAA ATT ATC AAT GTG	96
Ile Ser Val Ser Tyr Asn Pro Ala Thr Ala Lys Glu Ile Ile Asn Val	
20 25 30	
GGG CAT TCC TTC CAT GTA AAT TTT GAG GAC AAC GAT AAC CGA TCA GTG	144
Gly His Ser Phe His Val Asn Phe Glu Asp Asn Asp Asn Arg Ser Val	
35 40 45	
CTG AAA GGT GGT CCT TTC TCT GAC AGC TAC AGG CTC TTT CAG TTC CAT	192
Leu Lys Gly Gly Pro Phe Ser Asp Ser Tyr Arg Leu Phe Gln Phe His	
50 55 60	
TTT CAC TGG GGC AGT ACA AAT GAG CAT GGT TCA GAA CAT ACA GTG GAT	240
Phe His Trp Gly Ser Thr Asn Glu His Gly Ser Glu His Thr Val Asp	
65 70 75 80	
GGA GTC AAA TAT TCT GCC GAG CTT CAC GTG GCT CAC TGG AAT TCT GCA	288
Gly Val Lys Tyr Ser Ala Glu Leu His Val Ala His Trp Asn Ser Asp	
85 90 95	
AAG TAC TCC AGC CTT GCT GAA GCT GCC TCA AAG GCT GAT GGT TTG GCA	336
Lys Tyr Ser Ser Leu Ala Glu Ala Ala Ser Lys Ala Asp Gly Leu Ala	
100 105 110	
GTT ATT GGT GTT TTG ATG AAG GTT GGT GAG GCC AAC CCA AAG CTG CAG	384
Val Ile Gly Val Leu Met Lys Val Gly Glu Ala Asn Pro Lys Leu Gln	
115 120 125	
AAA GTA CTT GAT GCC CTC CAA GCA ATT AAA ACC AAG GGC AAA CGA GCC	432
Lys Val Leu Asp Ala Leu Gln Ala Ile Lys Thr Lys Gly Lys Arg Ala	
130 135 140	
CCA TTC ACA AAT TTT GAC CCC TCT ACT CTC CTT CCT TCA TCC CTG GAT	480
Pro Phe Thr Asn Phe Asp Pro Ser Thr Leu Leu Pro Ser Ser Leu Asp	
145 150 155 160	
TTC TGG ACC TAC CCT GGC TCT CTG ACT CAT CCT CCT CTT TAT GAG AGT	528
Phe Trp Thr Tyr Pro Gly Ser Leu Thr His Pro Pro Leu Tyr Glu Ser	
165 170 175	
GTA ACT TGG ATC ATC TGT AAG GAG AGC ATC AGT GTC AGT TCA GAG CAG	576
Val Thr Trp Ile Ile Cys Lys Glu Ser Ile Ser Val Ser Ser Glu Gln	
180 185 190	

TTG GCA CAA TTC CGG AGC CTT CTA TCA AT
 Leu Ala Gln Phe Arg Ser Leu Leu Ser
 195 200

605

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 201 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Val Asp Ile Lys Thr Ser Glu Thr Lys His Asp Thr Ser Leu Lys Pro
 1 5 10 15
 Ile Ser Val Ser Tyr Asn Pro Ala Thr Ala Lys Glu Ile Ile Asn Val
 20 25 30
 Gly His Ser Phe His Val Asn Phe Glu Asp Asn Asp Asn Arg Ser Val
 35 40 45
 Leu Lys Gly Gly Pro Phe Ser Asp Ser Tyr Arg Leu Phe Gln Phe His
 50 55 60
 Phe His Trp Gly Ser Thr Asn Glu His Gly Ser Glu His Thr Val Asp
 65 70 75 80
 Gly Val Lys Tyr Ser Ala Glu Leu His Val Ala His Trp Asn Ser Ala
 85 90 95
 Lys Tyr Ser Ser Leu Ala Glu Ala Ala Ser Lys Ala Asp Gly Leu Ala
 100 105 110
 Val Ile Gly Val Leu Met Lys Val Gly-Glu Ala Asn Pro Lys Leu Gln
 115 120 125
 Lys Val Leu Asp Ala Leu Gln Ala Ile Lys Thr Lys Gly Lys Arg Ala
 130 135 140
 Pro Phe Thr Asn Phe Asp Pro Ser Thr Leu Leu Pro Ser Ser Leu Asp
 145 150 155 160
 Phe Trp Thr Tyr Pro Gly Ser Leu Thr His Pro Pro Leu Tyr Glu Ser
 165 170 175
 Val Thr Trp Ile Ile Cys Lys Glu Ser Ile Ser Val Ser Ser Glu Gln
 180 185 190
 Leu Ala Gln Phe Arg Ser Leu Leu Ser
 195 200

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 559 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 2..469

(ix) FEATURE:

(A) NAME/KEY: mat_peptide
 (B) LOCATION: 2..469

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

C GGC TCC GGG CGG GCG TGG CCA GTG ACT AGA AGG CGA GGC GCC GCG 46
 Gly Ser Gly Arg Ala Trp Pro Val Thr Arg Arg Arg Gly Ala Ala
 1 5 10 15

GGA CCA TGG CGG CGG CGG ACG AGC GGA GTC CAG AGG CGA GAA GAC 94
 Gly Pro Trp Arg Arg Arg Thr Ser Gly Val Gln Arg Arg Gly Ala Asp
 20 25 30

GAG GAA GAG GAG GAG CAG TTG GTT CTG GTG GAA TTA TCA GGA ATT ATT 142
 Glu Glu Glu Glu Glu Gln Leu Val Leu Val Glu Leu Ser Gly Ile Ile
 35 40 45

GAT TCA GAC TTC CTC TCA AAA TGT GAA AAT AAA TGC AAG GTT TTG GGC 190
 Asp Ser Asp Phe Leu Ser Lys Cys Glu Asn Lys Cys Lys Val Leu Gly
 50 55 60

ATT GAC ACT GAG AGG CCC ATT CTG GCA ATG GAC AGC TGT GTC TTT GGT 238
 Ile Asp Thr Glu Arg Pro Ile Leu Ala Met Asp Ser Cys Val Phe Ala
 65 70 75

GGG GAG TAT GAA GAC ACT CTA GGG ACC TGT GTT ATA TTT GAA GAA AAT 286
 Gly Glu Tyr Glu Asp Thr Leu Gly Thr Cys Val Ile Phe Glu Glu Asn
 80 85 90 95

GTT GAA CAT GCT GAT ACA GAA GGC AAT AAT AAA ACA GTG CTA AAA TAT 334
 Val Glu His Ala Asp Thr Glu Gly Asn Asn Lys Thr Val Leu Lys Tyr
 100 105 110

AAA TGC CAT ACA ATG AAG AAG CTC AGC ATG ACA AGA ACT CTC CTG ACA 382
 Lys Cys His Thr Met Lys Lys Leu Ser Met Thr Arg Thr Leu Leu Thr
 115 120 125

GAG AAG AAG GAA GGA GAA GAA AAC ATA GGT GGG GTG GAA TGG CTG CAA 430
 Glu Lys Lys Glu Gly Glu Glu Asn Ile Gly Gly Val Glu Trp Leu Gln
 130 135 140

ATA AGG ATA TGG TTT CTC CCT TTG ACC CAA CAG GTT TGT TAACCTTTCT 479
 Ile Arg Ile Trp Phe Leu Pro Leu Thr Gln Gln Val Cys

ACCATGAAAT TGAGGACGAG GAAGTGGTAG CTTTCAGCCC CGTTAAATCT TTGGATTGG

539

GAGGGGGTGG GGTTCATG

559

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 156 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Gly Ser Gly Arg Ala Trp Pro Val Thr Arg Arg Arg Gly Ala Ala Gly
1 5 10 15
Pro Trp Arg Arg Arg Thr Ser Gly Val Gln Arg Arg Glu Asp Glu
20 25 30
Glu Glu Glu Glu Gln Leu Val Leu Val Glu Leu Ser Gly Ile Ile Asp
35 40 45
Ser Asp Phe Leu Ser Lys Cys Glu Asn Lys Cys Lys Val Leu Gly Ile
50 55 60
Asp Thr Glu Arg Pro Ile Leu Ala Met Asp Ser Cys Val Phe Ala Gly
65 70 75 80
Glu Tyr Glu Asp Thr Leu Gly Thr Cys Val Ile Phe Glu Glu Asn Val
85 90 95
Glu His Ala Asp Thr Glu Gly Asn Asn Lys Thr Val Leu Lys Tyr Lys
100 105 110
Cys His Thr Met Lys Lys Leu Ser Met Thr Arg Thr Leu Leu Thr Glu
115 120 125
Lys Lys Glu Gly Glu Glu Asn Ile Gly Gly Val Glu Trp Leu Gln Ile
130 135 140
Arg Ile Trp Phe Leu Pro Leu Thr Gln Gln Val Cys
145 150 155

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 409 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTGGCAGAAG AAAGATAGGT TGGAGACAAT TGATTGCTCG ATGATATAAA ATGTTAAGTA	60
CCATGAATGN ATGCTGTTAG GCTGGAATGC GCCAAGATAA AAGGTGGGGC ATGGCATCAA	120
AAGGTAGGTC AACATATTAA ATAATTCAT GTATTGAAAT ATCCAGAAAA TATATAGACA	180
GATCTATAGA GATAGAACT GGTCTGCCCC GACTAGGGG TTGTCTAAGG ATAAGGAGCT	240
TCTTTTTTGG ATGGTGAAAT AACCTAAAA ATATTGTGCC ATGTTTGCA CAACTTTGTG	300
GAATATATTA AAAACCGGTT AATTGTACTC ACTAAATGT CCTCCTTCTT AAATTTAAGC	360
TGTTTNCCTGG ACAAGAAAAA GGGAAAGNNA CCAAGGGGNA AAAAATTTT	409

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 600 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGCCCTGGGCT TTGGGGGGGT CCCAAACATG GTATGCAGAA ATGTGATGGT TACAGGTCAG	60
TACAACCTCA GTCTTAGAA CCCCTCCACA CTTGAGCTCT GCACCCACTT TCCTGTCATT	120
TATTTATATA GGAAGTAGT TTTTTTTAGT TCGAGAGCCT TTCGAAGCTT AATTTATATT	180
CTTCTCTTGT ACCTTTTTTC TAAAAATACC AAAGATATTA CACAAAGGTA AATTAATGTT	240
CTCTGTTTTA TGCTTTATCT GATGGAGGCA AATATCCTCT TATTGTTGAT CAAAGGGGGC	300
AAAAGAATTT AGAGGCAAAAT GAACAAGCGA TAGGCTATGG CAACCTGAGA AAGAGAAGTG	360
NTCCTTCCAT CGTAAATTTA GNAGNCCAAG TAGGTAATGG GAACCAAAGT TGTTACTTTT	420
TTCTAGTAGT TATTTTCCC TTTTNNTTT TTGIGTACC TCTTACAGNG NCCCAAAACT	480
CCATTCTCTT TAAAGGGGTT TTTATGGGG GCTTACTGCA GGTAAAAAT TGGGGNCCAC	540
CATTTTAAAG GGGGGGCTAC CAGAAAGGAG GGGGGTCCCC NTTNCNAAAA AAAAAAATTG	600

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ATGCTTCCGG CTCGTATG

18

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGGTTTTCCC AGTCACGAC

19